Partial answer key to homework for Chapter 5

**5.5**

a) All coffee produced by this production line (or company if the line is representative of the

company).

b)

(107.98, 112.02)

> ybar<-110

> s<-7.1

> alpha<-0.05

> n<-50

> #Interval

> lower<-ybar - qt(p = 1 - alpha/2, df = n-1) \* s / sqrt(n)

> upper<-ybar + qt(p = 1 - alpha/2, df = n-1) \* s / sqrt(n)

> data.frame(lower, upper)

lower upper

1 107.9822 112.0178

c) We are 95% confident that the average caffeine content is between 107.98 and 112.02 milligrams

**5.6**

a) The width of the interval will be increased.

> ybar<-110

> s<-7.1

> alpha<-0.01

> n<-50

> #Interval

> lower<-ybar - qt(p = 1 - alpha/2, df = n-1) \* s / sqrt(n)

> upper<-ybar + qt(p = 1 - alpha/2, df = n-1) \* s / sqrt(n)

> data.frame(lower, upper)

lower upper

1 107.3091 112.6909

This has a higher range than 95% confidence interval

b) The width of the interval will be decreased

> # b

> ybar<-110

> s<-7.1

> alpha<-0.05

> n<-100

> #Interval

> lower<-ybar - qt(p = 1 - alpha/2, df = n-1) \* s / sqrt(n)

> upper<-ybar + qt(p = 1 - alpha/2, df = n-1) \* s / sqrt(n)

> data.frame(lower, upper)

lower upper

1 108.5912 111.4088

This has a lower range than 95% confidence interval

5.9

a)

(2.62, 2.98)

> ybar<-2.8

> s<-1.3

> alpha<-0.05

> n<-200

> #Interval

> lower<-ybar - qt(p = 1 - alpha/2, df = n-1) \* s / sqrt(n)

> upper<-ybar + qt(p = 1 - alpha/2, df = n-1) \* s / sqrt(n)

> data.frame(lower, upper)

lower upper

1 2.61873 2.98127

Perhaps this is a reasonable suggestion because the lower bound of 95% confidence interval is greater than 2.

b)

As the records were only from the courthouse of a single city, the confidence interval only really

applies to two-time offenders from that city.

c)

It would be unadvisable to generalize the results to a larger population other than two-time

offenders from that city.

5.10

a)

(8.54, 9.50)

> #5.10

> ybar<-9.02

> s<-1.12

> alpha<-0.1

> n<-40

>

> #Interval

> lower<-ybar - qt(p = 1 - alpha/2, df = n-1) \* s / sqrt(n)

> upper<-ybar + qt(p = 1 - alpha/2, df = n-1) \* s / sqrt(n)

> data.frame(lower, upper)

lower upper

1 8.72163 9.31837

b)

Assuming the orange trees used in the sample are representative of all orange trees (the problem

gives no reason to think otherwise), the population is all orange trees.

5.36

a) Incoming freshman for the school district

b) The 95% confidence interval on the average reading speed for all incoming freshman is (7.9, 10.3). The 95% confidence interval on the average reading comprehension for all incoming freshman is (77.0, 87.1).

> ex5.36<-read.table("C:/sdata/ex5-36.txt", header=T, row.names=1, sep=",")

> nrow(ex5.36)

[1] 20

> head(ex5.36)

ReadTime Comprehension

1 5 60

2 7 76

3 15 76

4 12 90

5 8 81

6 7 75

> t.test(x = ex5.36[,1], conf.level = 0.95)

One Sample t-test

data: ex5.36[, 1]

t = 15.816, df = 19, p-value = 2.16e-12

alternative hypothesis: true mean is not equal to 0

95 percent confidence interval:

7.895733 10.304267

sample estimates:

mean of x

9.1

> t.test(x = ex5.36[,2,], conf.level = 0.95)

One Sample t-test

data: ex5.36[, 2, ]

t = 33.727, df = 19, p-value < 2.2e-16

alternative hypothesis: true mean is not equal to 0

95 percent confidence interval:

76.95819 87.14181

sample estimates:

mean of x

82.05

c)

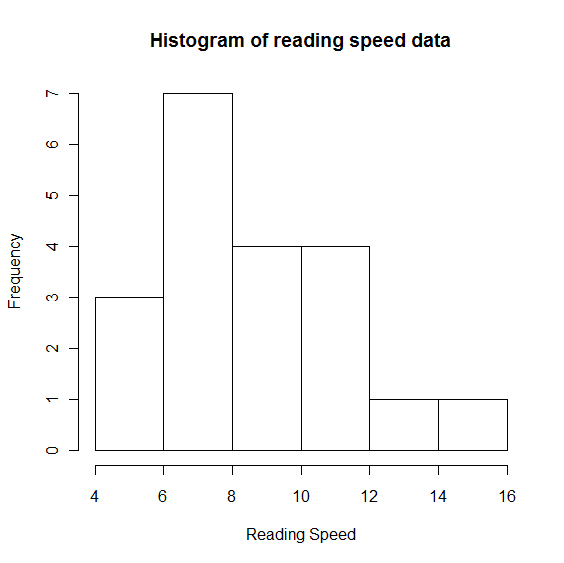
spd <- as.numeric(ex5.36[,1])

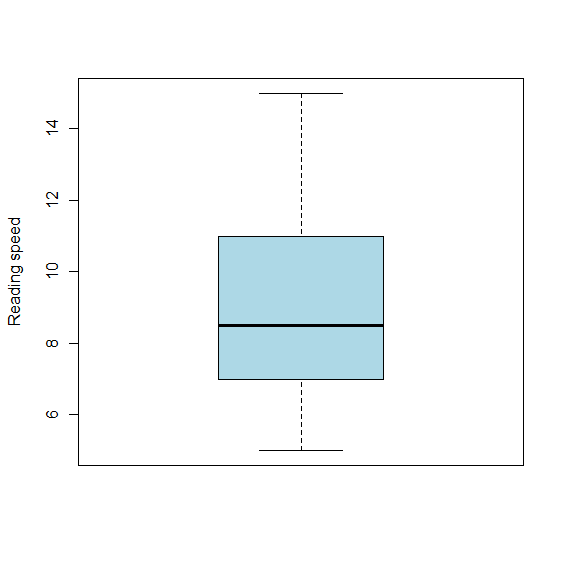
hist(spd, xlab = "Reading Speed", main = "Histogram of reading speed data")

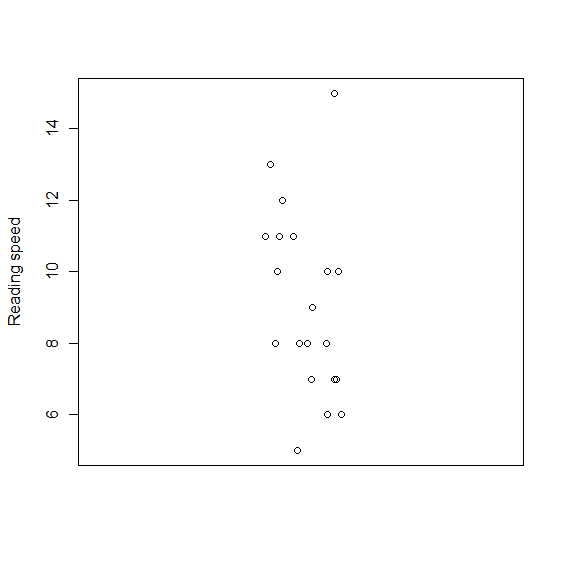
boxplot(x = spd, col = "lightblue", main = "", ylab = "Reading speed")

stripchart(x = spd, method = "jitter", vertical = TRUE, pch = 1,

main = "", ylab = "Reading speed")







d) We are 95% confident that the mean reading speed for the population is between 7.9 and 10.3

minutes

Yes, because the 95% confidence interval on the average reading comprehension for all incoming freshman is all above 60.

5.39

(27592, 35341)

> #5.39

> #remember to change the file path when you use the code

> ex5.39<-read.table("C:\\ sdata\\ex5-39.txt", header=T, row.names=1, sep=",")

> nrow(ex5.39)

[1] 15

> head(ex5.39)

Miles

1 25

2 27

3 35

4 42

5 28

6 37

> t.test(x = ex5.39, conf.level = 0.99)

One Sample t-test

data: ex5.39

t = 24.177, df = 14, p-value = 8.109e-13

alternative hypothesis: true mean is not equal to 0

99 percent confidence interval:

27.59223 35.34110

sample estimates:

mean of x

31.46667

b)There is not evidence that the manufacturer’s claim is wrong because the confidence interval contains 35000 and larger values. However, we see that 35000 is near the upper bound. For this reason, a formal hypothesis test performed later in this section may conclude that there is “marginal evidence” that the claim is incorrect.